

Описание работы фермы LibreOffice для конвертации odt ods в различные форматы

BIV Group

1.0, : Final

Table of Contents

Описание работы с фермой LibreOffice	1
Введение	1
Архитектура	1
Докер контейнер с фермой LibreOffice	2
Использование докер контейнера	2
Нагрузочное тестирование	3

Описание работы с фермой LibreOffice

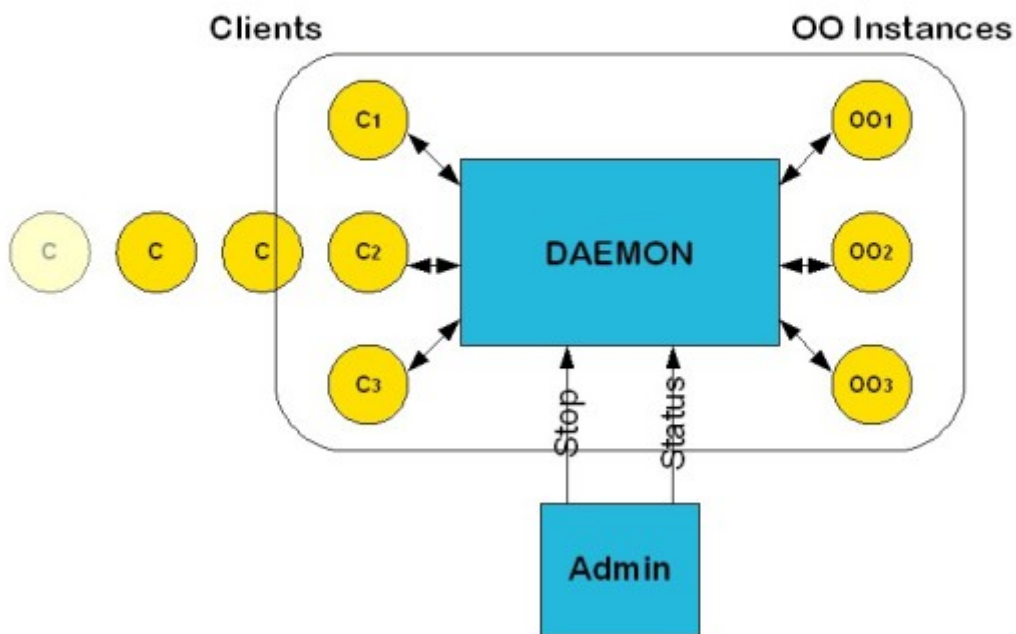
Введение

Ферма libreoffice предназначена для конвертации документов из одного формата в другой с помощью офисного пакета LibreOffice. Поддерживаемые способы конвертации:

1. odt → pdf
2. odt → docx (doc)
3. ods → xlsx

Архитектура

HOW IT WORKS



Ферма libreoffice представляет собой программу на java, которая:

1. Внешне прикидывается сервером libreoffice на сетевом порту (обычно 8100)
2. Принимает команды для libreoffice
3. Запускает заданное количество экземпляров libreoffice в виде сетевого сервера.
4. Следит за работоспособностью экземпляров libreoffice и при наличии проблем перезапускает экземпляр libreoffice. После заданного количества отработанных запросов перезапускает во избежание проблем экземпляр libreoffice.
5. Работает как балансировщик для запросов libreoffice. Выбирает наименее загруженный экземпляр и передает запрос на него.
6. Перед завершением работы убивает все экземпляры libreoffice

Докер контейнер с фермой LibreOffice

Для удобства развертывания фермы LibreOffice она оформлена как докер контейнер. Иерархия состоит из трех контейнеров:

1. `graal-mvn` - контейнер содержит oracle linux с установленным `graalvm`, `apache maven` и меркуриал.
2. `graal-mvn-libre` - контейнер расширяет предыдущий добавлением `libreoffice 6.2`
3. `libreferma` - многоэтапный контейнер, который при сборке (`docker build`) на первом шаге компилирует ферму из исходных кодов, а на втором шаге расширяет контейнер `graal-mvn-libre`, добавляя туда скомпилированную `jar` фермы, файлы конфигурации и скрипты запуска.

Oracle linux использован по причине того, что фирма Oracle выпускает новые версии `Graalvm` в виде контейнера исключительно на базе Oracle linux. Поэтому для легкости смены версии `Graalvm` удобно использовать контейнеры от Oracle.

ВНИМАНИЕ!!! . Если вы хотите поменять версию `libreoffice`, то необходимо помимо контейнера `graal-mvn-libre`, где надо поменять версию `libreoffice` и пересобрать, поменять версию `libreoffice` в скриптах и конфигурации фермы, так как там присутствуют пути на установленный `libreoffice`.

1. По-умолчанию контейнер запускает 10 экземпляров `libreoffice`. Если вы хотите поменять это количество, то необходимо поменять файл конфигурации, изменив пути до профилей пользователей `libreoffice` до нужного вам количества. Сами профили ферма создаст при запуске контейнера самостоятельно
2. Если вы хотите запустить контейнер в докер репозиторий, то необходимо поменять имена и ссылки на имена у всех трех контейнеров и последовательно их запустить, начиная с `graal-mvn`.

Использование докер контейнера

Докер контейнер выставляет наружу:

1. Порт 8100 для подключения к серверу. По нему передается путь до файлов для конвертации.
2. Каталог `/upload` где должны лежать файлы с отчетами и результаты конвертации.

ВНИМАНИЕ!!!

1. Контейнер предназначен для работы на том же самом компьютере или виртуалке, где работает вызывающий данных контейнер клиент. По сети контейнер вызвать нельзя!
2. Так как через сетевое соединение на порту 8100 передается путь до файла для конвертации, то клиент должен хранить файлы для конвертации по тому же пути (`/upload`), который используется контейнером, иначе пути не совпадут и ферма не найдет файл для конвертации! То есть, если у вас клиент работает внутри контейнера, то путь (`/upload`) должен быть смалпен на тот же `Volume`, что и в контейнере ферме

(/upload). Если клиент просто работает на виртуалке, то Volume должен быть сматлен на тот же каталог (/upload).

При необходимости каталог можно поменять. Для этого необходимо изменить путь в Dockerfile и пересобрать контейнер ферму.

Нагрузочное тестирование

Результаты нагрузочного тестирования <http://docs.bivgroup.com/display/techteam/LibreOffice>